EYE DETECTION USING WAVELETS AND ANN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology

In

Electronics & Instrumentation Engineering

By

DIPTI RANJAN PANDA

And

CHITTARANJAN NAYAK

Department of Electronics & Instrumentation Engineering

National Institute of Technology

Rourkela-769008

2007

EYE DETECTION USING WAVELETS AND ANN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology

In

Electronics & Instrumentation Engineering

By

DIPTI RANJAN PANDA

And

CHITTARANJAN NAYAK

Under the Guidance of

Prof. G.S.Rath

Department of Electronics & Instrumentation Engineering

National Institute of Technology

Rourkela-769008

2007

National Institute of Technology

Rourkela

CERTIFICATE

This is to certify that the thesis entitled, "Eye detection using wavelets and ANN" submitted by Sri Dipti ranjan Panda and Sri Chittaranjan Nayak in partial fulfillments for the requirements for the award of Bachelor of Technology Degree in Electronics & Instrumentation Engineering at National Institute of Technology, Rourkela (Deemed University) is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University / Institute for the award of any Degree or Diploma.

Date:

Prof. G. S. Rath
Dept. of Electronics & Instrumentation Engg
National Institute of Technology
Rourkela - 769008

ACKNOWLEDGEMENT

We would like to articulate our deep gratitude to our project guide Prof. G.S.Rath who has always been our motivation for carrying out the project. It is our pleasure to refer Microsoft Word exclusive of which the compilation of this report would have been impossible. An assemblage of this nature could never have been attempted with out reference to and inspiration from the works of others whose details are mentioned in reference section. We acknowledge out indebtedness to all of them. Last but not the least, our sincere thanks to all of our friends who have patiently extended all sorts of help for accomplishing this undertaking.

DIPTI RANJAN PANDA

CHITTARANJAN NAYAK

CONTENTS

ABSTRACT

A Biometric system provides perfect identification of individual based on a unique biological feature or characteristic possessed by a person such as finger print, hand writing, heart beat, face recognition and eye detection. Among them eye detection is a better approach since Human Eye does not change throughout the life of an individual. It is regarded as the most reliable and accurate biometric identification system available.

In our project we are going to develop a system for ‘eye detection using wavelets and ANN’ with software simulation package such as matlab 7.0 tool box in order to verify the uniqueness of the human eyes and its performance as a biometric. Eye detection involves first extracting the eye from a digital face image, and then encoding the unique patterns of the eye in such a way that they can be compared with pre-registered eye patterns. The eye detection system consists of an automatic segmentation system that is based on the wavelet transform, and then the Wavelet analysis is used as a pre-processor for a back propagation neural network with conjugate gradient learning. The inputs to the neural network are the wavelet maxima neighborhood coefficients of face images at a particular scale. The output of the neural network is the classification of the input into an eye or non-eye region. An accuracy of 81% is observed for test images under different environment conditions not included during training.

Eye detection system is being extensively used in biometrics security solutions by U.S. Department of Defense (DOD), which includes access control to physical facilities, security systems or information databases. Suspect tracking, surveillance and intrusion detection and by various Intelligence agencies through out the world, also in the corrections/laws enforcement marketplaces.

Chapter 1

GENERAL INTRODUCTION

1.1. Biometric Technology:

A biometric system provides automatic recognition of an individual based on some sort of unique feature or characteristic possessed by the individual. Biometric systems have been developed based on fingerprints, facial features, voice, hand geometry, handwriting, the retina, and the one presented in this project, the eye.

Biometric systems work by first capturing a sample of the feature, such as recording a digital sound signal for voice recognition, or taking a digital color image for eye detection. The sample is then transformed using some sort of mathematical function into a biometric template. The biometric template will provide a normalized, efficient and highly discriminating representation of the feature, which can then be objectively compared with other templates in order to determine identity. Most biometric systems allow two modes of operation. A training mode or enrolment mode for adding templates to a database, and an identification mode, where a template is created for an individual and then a match is searched for in the database of pre-enrolled templates.

A good biometric is characterized by use of a feature that is; highly unique so that the chance of any two people having the same characteristic will be minimal, stable so that the feature does not change over time, and be easily captured in order to provide convenience to the user, and prevent misrepresentation of the feature.

1.2. EYE: The Perfect ID

The randomness and uniqueness of human eye patterns is a major breakthrough in the search for quicker, easier and highly reliable forms of automatic human identification, where the human eye serves as a type of 'biological passport, PIN or password'.

Results of a study by John Daugman and Cathryn, of over two million different pairs of human eyes in images taken from volunteers in Britain, USA and Japan show that no two eye patterns were the same in even as much as one-third of their form. Even genetically identical faces - for example from twins or in the probable future, from human clones - have different eye patterns.

The implications of eye detection are highly significant at a time when organizations such as banks and airlines are looking for more effective security measures. The possible applications of eye detection span all aspects of daily life, from computer login, national border controls and secure access to bank cash machine accounts, to ticket-less air travel, access to premises such as the home and office, benefits entitlement and credit card authentication.

Compared with other biometric technologies, such as face, speech and finger recognition, eye recognition can easily be considered as the most reliable form of biometric. However, there have been no independent trials of the technology, and source code for systems is not available in working condition.

1.3. Objective:

The objective will be to implement an open-source eye detection system in order to verify the claimed performance of the technology. This project is based on a novel method, which is robust and efficient in extracting eye windows using Wavelets and Neural Networks. Wavelet analysis is used as a pre processor for a back propagation neural network with conjugate gradient learning. The inputs to the neural network are the wavelet maxima neighborhood coefficients of face images at a particular scale. The output of the neural network is the classification of the input into an eye or non-eye region. The updated weight and bias values for a particular person is stored in a database. The image to be verified is wavelet transformed before being applied to the neural network with those updated weight and bias values. The person is identified when the neural network output of one of the test images matches with that of the verified image. . An accuracy of 90% is observed for test images under different environment conditions not included during training.

Chapter 2

FUNDAMENTALS OF WAVELET TRANSFORM

2.1. INTRODUCTION

The transform of a signal is just another form of representing the signal. It does not change the information content present in the signal. The Wavelet Transform provides a time-frequency representation of the signal. It was developed to overcome the shortcoming of the Short Time Fourier Transform (STFT), which can also be used to analyze non-stationary signals. While STFT gives a constant resolution at all frequencies, the Wavelet Transform uses multi-resolution technique by which different frequencies are analyzed with different resolutions.

A wave is an oscillating function of time or space and is periodic. In contrast, wavelets are localized waves. They have their energy concentrated in time or space and are suited to analysis of transient signals. While Fourier Transform and STFT use waves to analyze signals, the Wavelet Transform uses wavelets of finite energy.

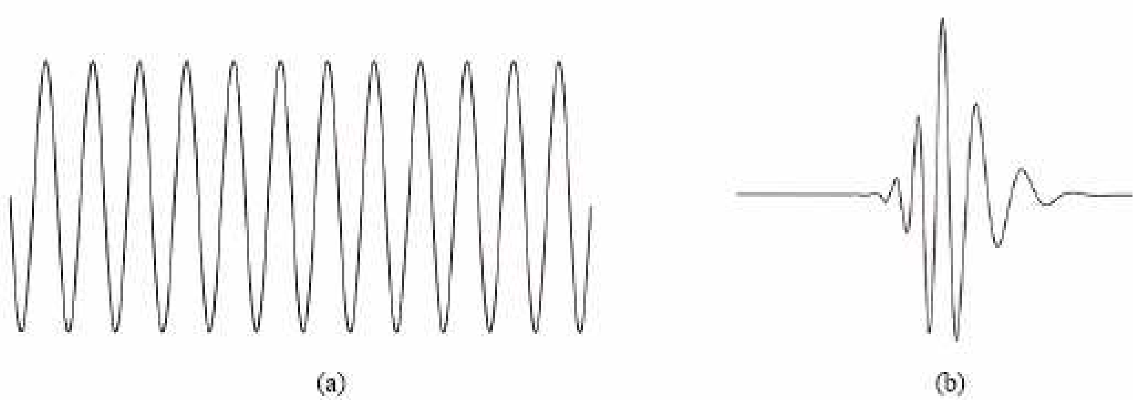


Figure 2.1 Demonstration of (a) a Wave and (b) a Wavelet [2].

The wavelet analysis is done similar to the STFT analysis. The signal to be analyzed is multiplied with a wavelet function just as it is multiplied with a window function in STFT, and then the transform is computed for each segment generated.

However, unlike STFT, in Wavelet Transform, the width of the wavelet function changes with each spectral component. The Wavelet Transform, at high frequencies, gives good time resolution and poor frequency resolution, while at low frequencies; the Wavelet Transform gives good frequency resolution and poor time resolution.

2.2 The Continuous Wavelet Transform and the Wavelet Series

The Continuous Wavelet Transform (CWT) is provided by equation 2.1, where x(t) is the signal to be analyzed. ψ(t) is the mother wavelet or the basis function. All the wavelet functions used in the transformation are derived from the mother wavelet through translation (shifting) and scaling (dilation or compression).

$$X_{WT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \cdot \psi^{*}\left(\frac{t-\tau}{s}\right) dt \quad \text{............2.1}$$

The mother wavelet used to generate all the basis functions is designed based on some desired characteristics associated with that function. The translation parameter τ relates to the location of the wavelet function as it is shifted through the signal. Thus, it corresponds to the time information in the Wavelet Transform. The scale parameter s is defined as |1/frequency| and corresponds to frequency information. Scaling either dilates (expands) or compresses a signal. Large scales (low frequencies) dilate the signal and provide detailed information hidden in the signal, while small scales (high frequencies) compress the signal and provide global information about the signal. Notice that the Wavelet Transform merely performs the convolution operation of the signal and the basis function. The above analysis becomes very useful as in most practical applications; high frequencies (low scales) do not last for a long duration, but instead, appear as short bursts, while low frequencies (high scales) usually last for entire duration of the signal.

The Wavelet Series is obtained by discretizing CWT. This aids in computation of CWT using computers and is obtained by sampling the time-scale plane. The sampling rate can be changed accordingly with scale change without violating the Nyquist criterion. Nyquist criterion states that, the minimum sampling rate that allows reconstruction of the original signal is 2ω radians, where ω is the highest frequency in the signal. Therefore, as the scale goes higher (lower frequencies), the sampling rate can be decreased thus reducing the number of computations.

The Wavelet Series is obtained by discretizing CWT. This aids in computation of CWT using computers and is obtained by sampling the time-scale plane. The sampling rate can be changed accordingly with scale change without violating the Nyquist criterion. Nyquist criterion states that, the minimum sampling rate that allows reconstruction of the original signal is 2ω radians, where ω is the highest frequency in the signal. Therefore, as the scale goes higher (lower frequencies), the sampling rate can be decreased thus reducing the number of computations.

Chapter 3

DISCRETE WAVELET TRANSFORM

3.1. INTRODUCTION

The Wavelet Series is just a sampled version of CWT and its computation may consume significant amount of time and resources, depending on the resolution required. The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found to yield a fast computation of Wavelet Transform. It is easy to implement and reduces the computation time and resources required.

The foundations of DWT go back to 1976 when techniques to decompose discrete time signals were devised. Similar work was done in speech signal coding which was named as sub-band coding. In 1983, a technique similar to sub-band coding was developed which was named pyramidal coding. Later many improvements were made to these coding schemes, which resulted in efficient multi-resolution analysis schemes.

In CWT, the signals are analyzed using a set of basis functions, which relate to each other by simple scaling and translation. In the case of DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales.

3.2. DWT and Filter Banks

3.2.1 Multi-Resolution Analysis using Filter Banks

Filters are one of the most widely used signal processing functions. Wavelets can be realized by iteration of filters with rescaling. The resolution of the signal, which is a measure of the amount of detail information in the signal, is determined by the filtering operations, and the scale is determined by upsampling and downsampling (subsampling) operations.

The DWT is computed by successive lowpass and highpass filtering of the discrete time-domain signal as shown in figure 3.1. This is called the Mallat algorithm or Mallat-tree decomposition. Its significance is in the manner it connects the continuous-time multiresolution to discrete-time filters. In the figure, the signal is denoted by the sequence x[n], where n is an integer. The low pass filter is denoted by G_0 while the high pass filter is denoted by H_0. At each level, the high pass filter produces detail information; d[n], while the low pass filter associated with scaling function produces coarse approximations, a[n].

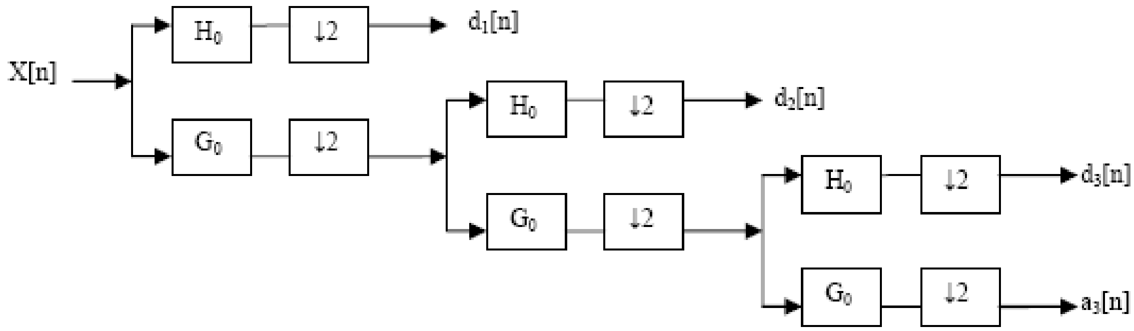


Figure 3.1 Three-level wavelet decomposition tree.

At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. In accordance with Nyquist's rule if the original signal has a highest frequency of ω, which requires a sampling frequency of 2ω radians, then it now has a highest frequency of ω/2 radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale.

With this approach, the time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition.

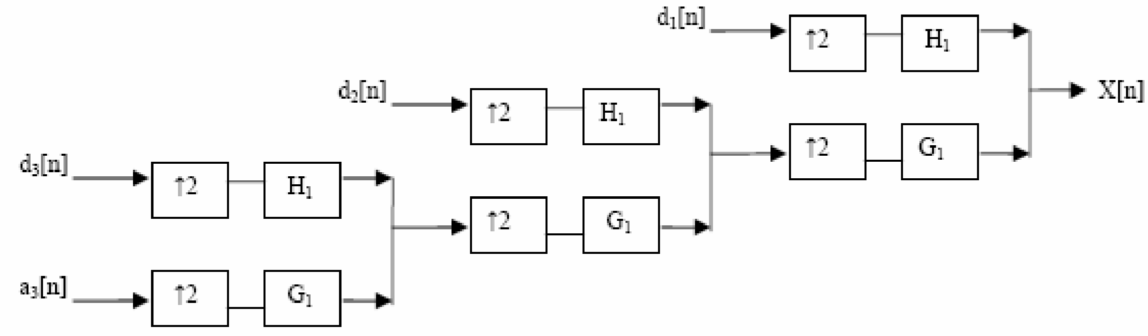


Figure 3.2 Three-level wavelet reconstruction tree.

Figure 3.2 shows the reconstruction of the original signal from the wavelet coefficients. Basically, the reconstruction is the reverse process of decomposition. The approximation and detail coefficients at every level are upsampled by two, passed through the low pass and high pass synthesis filters and then added. This process is continued through the same number of levels as in the decomposition process to obtain the original signal. The Mallat algorithm works equally well if the analysis filters, G_0 and H_0, are exchanged with the synthesis filters, G_1 & H_1.

3.2.2 Conditions for Perfect Reconstruction

In most Wavelet Transform applications, it is required that the original signal be synthesized from the wavelet coefficients. To achieve perfect reconstruction the analysis and synthesis filters have to satisfy certain conditions. Let $G_0(z)$ and $G_1(z)$ be the low pass analysis and synthesis filters, respectively and $H_0(z)$ and $H_1(z)$ the high pass analysis and synthesis filters respectively. Then the filters have to satisfy the following two conditions:

$$G_0(-z)\, G_1(z) + H_0(-z).\, H_1(z) = 0 \qquad (3.1)$$

$$G_0(z)\, G_1(z) + H_0(z).\, H_1(z) = 2z^{-d} \qquad (3.2)$$

The first condition implies that the reconstruction is aliasing-free and the second condition implies that the amplitude distortion has amplitude of one. It can be observed that the perfect reconstruction condition does not change if we switch the analysis and synthesis filters.

There are a number of filters, which satisfy these conditions. But not all of them give accurate Wavelet Transforms, especially when the filter coefficients are quantized. The accuracy of the Wavelet Transform can be determined after reconstruction by calculating the Signal to Noise Ratio (SNR) of the signal. Some applications like pattern recognition do not need reconstruction, and in such applications, the above conditions need not apply.

3.2.3 Classification of wavelets

We can classify wavelets into two classes: (a) orthogonal and (b) biorthogonal. Based on the application, either of them can be used.

(a) Features of orthogonal wavelet filter banks

The coefficients of orthogonal filters are real numbers. The filters are of the same length and are not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other by

$$H_0(z) = z^{-N} G_0(-z^{-1}) \qquad \text{(3.3)}$$

The two filters are alternated flip of each other. The alternating flip automatically gives double-shift orthogonality between the lowpass and highpass filters, i.e., the scalar product of the filters, for a shift by two is zero. i.e., $\Sigma G[k]\, H[k-2l] = 0$, where $k,l \in Z$. Filters that satisfy equation 3.3 are known as Conjugate Mirror Filters (CMF). Perfect reconstruction is possible with alternating flip.

Also, for perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Orthogonal filters offer a high number of vanishing moments. This property is useful in many signal and image processing applications. They have regular structure, which leads to easy implementation and scalable architecture.

(b)Features of biorthogonal wavelet filter banks

In the case of the biorthogonal wavelet filters, the low pass and the high pass filters do not have the same length. The low pass filter is always symmetric, while the high pass filter could be either symmetric or anti-symmetric. The coefficients of the filters are either real numbers or integers.

For perfect reconstruction, biorthogonal filter bank has all odd length or all even length filters. The two analysis filters can be symmetric with odd length or one symmetric and the other antisymmetric with even length. Also, the two sets of analysis and synthesis filters must be dual. The linear phase biorthogonal filters are the most popular filters for data compression applications.

3.3 Wavelet Families

There are a number of basis functions that can be used as the mother wavelet for Wavelet Transformation. Since the mother wavelet produces all wavelet functions used in the transformation through translation and scaling, it determines the characteristics of the resulting Wavelet Transform. Therefore, the details of the particular application should be taken into account and the appropriate mother wavelet should be chosen in order to use the Wavelet Transform effectively.

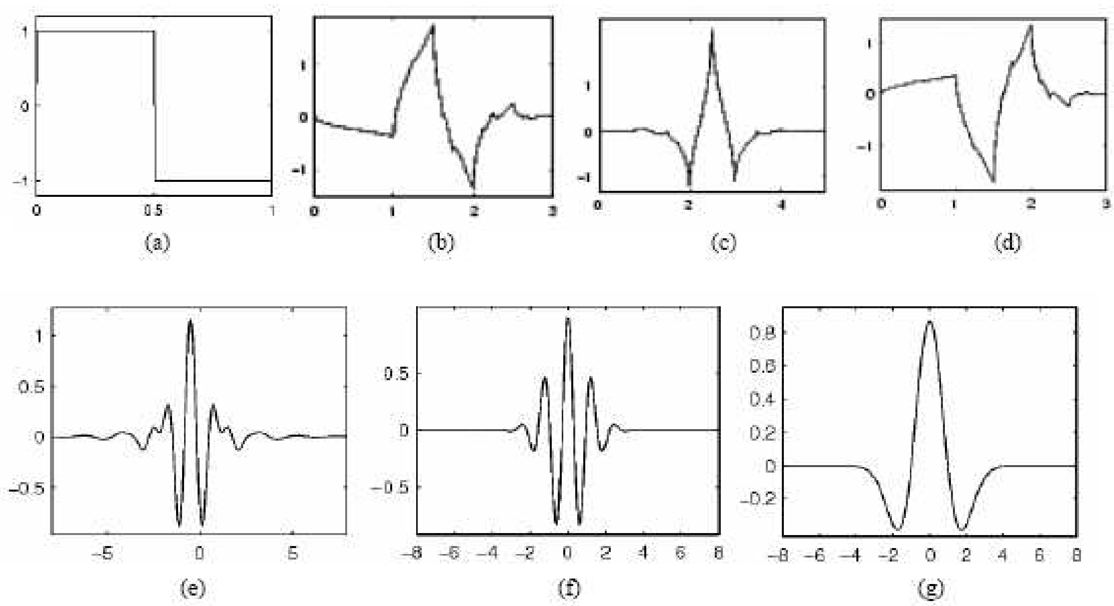


Figure 3.3 Wavelet families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat.

Figure 3.3 illustrates some of the commonly used wavelet functions. Haar wavelet is one of the oldest and simplest wavelet. Therefore, any discussion of wavelets starts with the Haar wavelet. Daubechies wavelets are the most popular wavelets. They represent the foundations of wavelet signal processing and are used in numerous applications. These are also called Maxflat wavelets as their frequency responses have maximum flatness at frequencies 0 and π. This is a very desirable property in some applications. The Haar, Daubechies, Symlets and Coiflets are compactly supported orthogonal wavelets. These wavelets along with Meyer wavelets are capable of perfect reconstruction. The Meyer, Morlet and Mexican Hat wavelets are symmetric in shape. The wavelets are chosen based on their shape and their ability to analyze the signal in a particular application.

Chapter 4

METHODS OF EYE DETECTION

4.1 INTRODUCTION:

A lot of research work has been published in the field of eye detection in the last decade. Various techniques have been proposed using texture, depth, shape and color information or combinations of these for eye detection. Vezhnevets focus on several landmark points (eye corners, iris border points), from which the approximate eyelid contours are estimated. The upper eyelid points are found using on the observation that eye border pixels are significantly darker than surrounding skin and sclera. The detected eye boundary points are filtered to remove outliers and a polynomial curve is fitted to the remaining boundary points. The lower lid is estimated from the known iris and eye. Some of the famous eye detection techniques are discussed below.

4.2 TEMPLATE MATCHING METHOD:

Reinders present a method where based on the technique of template matching the positions of the eyes on the face image can be followed throughout a sequence of video images. Template matching is one of the most typical techniques for feature extraction. Correlation is commonly exploited to measure the similarity between a stored template and the window image under consideration. Templates should be deliberately designed to cover variety of possible image variations. During the search in the whole image, scale and rotation should also be considered carefully to speed up the process. To increase the robustness of the tracking scheme the method automatically generates a codebook of images representing the encountered different appearances of the eyes. Yuille first proposed using deformable templates in locating human eye. The weaknesses of the deformable templates are that the processing time is lengthy and success relies on the initial position of the template. Lam introduced the concept of eye corners to improve the deformable template approach.

4.3 USING PROJECTION FUNCTION:

Saber and Jeng proposed to use facial features geometrical structure to estimate the location of eyes. Takacs developed iconic filter banks for detecting facial landmarks. projection functions have also been employed to locate eye windows. Feng and Yeun developed a variance projection function for locating the corner points of the eye. Zhou and Geng propose a hybrid projection function to locate the eyes. By combining an integral projection function, which considers mean of intensity, and a variance projection function, which considers the variance of intensity, the hybrid function better captures the vertical variation in intensity of the eyes. Kumar suggest a technique in which possible eye areas are localized using a simple thresholding in color space followed by a connected component analysis to quantify spatially connected regions and further reduce the search space to determine the contending eye pair windows. Finally the mean and variance projection functions are utilized in each eye pair window to validate the presence of the eye. Feng and Yeun employ multi cues for eye detection on gray images using variance projection function.

4.4 IR METHOD:

The most common approach employed to achieve eye detection in real-time is by using infrared lighting to capture the physiological properties of eyes and an appearance-based model to represent the eye patterns. The appearance-based approach detects eyes based on the intensity distribution of the eyes by exploiting the differences in appearance of eyes from the rest of the face. This method requires a significant number of training data to enumerate all possible appearances of eyes i.e. representing the eyes of different subjects, under different face orientations, and different illumination conditions. The collected data is used to train a classifier such as a neural net or support vector machine to achieve detection.

4.5 SUPPORT VECTOR MACHINES (SVMs).

Support Vector Machines (SVMs) have been recently proposed by Vapnik and his co-workers as a very effective method for general-purpose pattern recognition. Intuitively, given a set of points belonging to two classes, a SVM finds the hyper-plane that separates the largest possible fraction of points of the same class to the same side while maximizing the distances from either class to the hyper-plane. This hyper-plane is called Optimal Separating Hyper-plane (OSH). It minimizes the risk of misclassifying not only the samples in the training set but also the unseen samples in the test set. The application of SVMs to computer vision area has emerged recently. Osuna train a SVM for face detection, where the discrimination is between two classes: face and non-face, each with thousands of samples. Guo and Stan show that the SVMs can be effectively trained for face recognition and is a better learning algorithm than the nearest center approach.

Graph Matching. After all images, including the gallery images and the probe images, are extracted using EBGM procedure, the faces are represented as labelled face graphs. The matching procedure then involves the distance computation of the jets between different graphs, which is represented as:

$$S_a(J, J') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}$$

On the FERET dataset, the algorithm performs impressively well for the frontal images with recognition accuracy of 98%. For half rotated and profile images, the performance degrades to 57% and 84%, respectively; however, since these are difficult cases in face detection and recognition systems, the results are still comparatively good.

4.6 Hidden Markov Models (HMMs):

HMMs are generally used for the statistical modelling of non-stationary vector time series. By considering the facial configurable information as a time varying sequence, HMMs can be applied to face recognition. The most significant facial features of a frontal face image, including the hair, forehead, eyes, nose and mouth, occur in a natural order from top to bottom, even if the image has small rotations in the image plane, and/or rotations in the plane perpendicular to the image plane. Based on this observation, the image of a face may be modeled using a one-dimensional HMM by assigning each of these regions a state as illustrated in Figure 4.1.

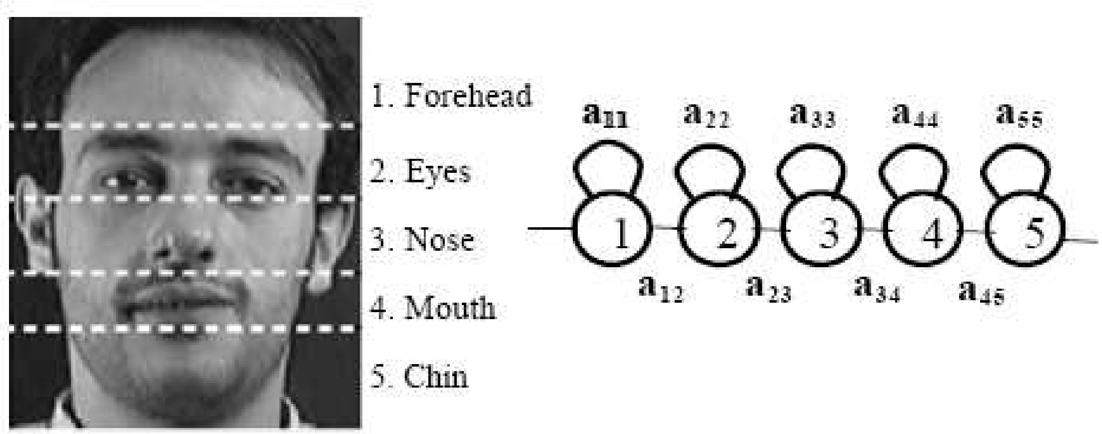


Figure 4.1. A Top-to-Bottom 5 states HMM.

Given a face image for one subject in the training set, the goal of the training stage is to optimize the parameters to best describe the observation. Recognition is carried out by matching the test image against each of the trained models. To complete this procedure, the image is converted to an observation sequence and the likelihood is computed for each stored model. The model with the highest likelihood reveals the identity of the unknown face. The HMM approach has shown the ability to yield satisfactory recognition rates. However, HMMs are processor intensive models, which implies that the algorithm may run slowly. The HMM lead to the efficient detection of eye strips.

4.7 WAVELET BASED METHOD:

Our project is based on this method of eye detection. Wavelet decomposition provides local information in both space domain and frequency domain. Despite the equal subband sizes, different subbands carry different amounts of information. The letter ‘L’ stands for low frequency and the letter ‘H’ stands for high frequency. The left upper band is called LL band because it contains low frequency information in both the row and column directions. The LL band is a coarser approximation to the original image containing the overall information about the whole image. The LH subband is the result of applying the filter bank column wise and extracts the facial features very well. The HL subband, which is the result of applying the filter bank row wise, extracts the outline of the face boundary very well. While the HH band shows the high frequency component of the image in non-horizontal, non-vertical directions it proved to be a redundant subband and was not considered having significant information about the face. This observation was made at all resolutions of the image. This is the first level decomposition. Finally a fixed no. of maximum peaks are selected from LH subband, which are fed as inputs to the neural network back propagation model or RBF or neuro-fuzzy model is used to train that required network. According to the outputs of those peaks, after being passed through the updated weight and bias values, they are categorized into eye parts and non-eye parts. our project is based on this method of eye detection.

Chapter 5

WAVELET BASED METHOD FOR EYE DETECTION

5.1 INTRODUCTION:

The system consists mainly of two stages training and detection stage. A block diagram of these two stages is shown in Figure 1.

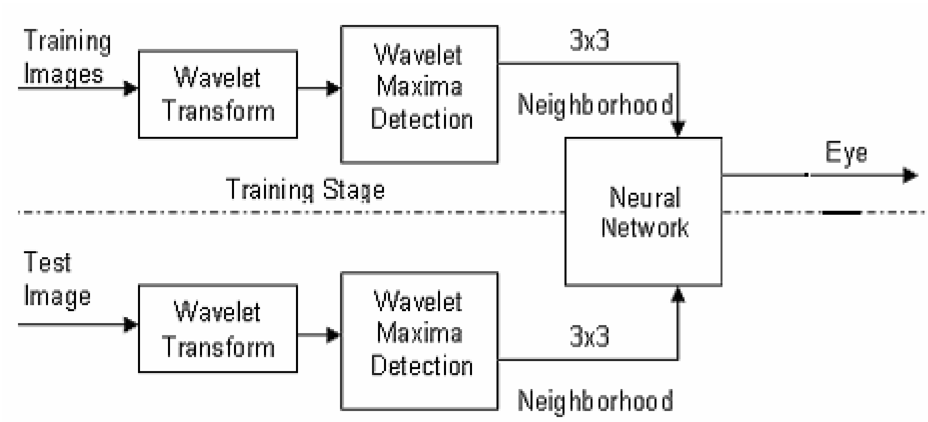


Figure 5.1 - Block Diagram

5.2 Acquisition of Training Data:

The training data typically consists of 50 images of different persons with different hairstyles, different illumination conditions and varying facial expressions. Some of the images have different states of the eye such as eyes closed. The size of the images varies from 64x64 to 256x256.

5.3 Discrete Wavelet Transform:

Wavelet decomposition provides local information in both space domain and frequency domain. Despite the equal subband sizes, different subbands carry different amounts of information. The letter 'L' stands for low frequency and the letter 'H' stands for high frequency. The left upper band is called LL band because it contains low frequency information in both the row and column directions. The LL band is a coarser approximation to the original image containing the overall information about the whole image. The LH subband is the result of applying the filter bank column wise and extracts the facial features very well. The HL subband, which is the result of applying the filter bank row wise, extracts the outline of the face boundary very well. While the HH band shows the high frequency component of the image in non-horizontal, non-vertical directions it proved to be a redundant subband and was not considered having significant information about the face. This observation was made at all resolutions of the image. This is the first level decomposition. A CDF (2, 2) biorthogonal wavelet is used. Gabor Wavelets seem to be the most probable candidate for feature extraction. But they suffer from certain limitations i.e. they cannot be implemented using Lifting Scheme and secondly the Gabor Wavelets form a non-orthogonal set thus making the computation of wavelet coefficients difficult and expensive. Special hardware is required to make the algorithm work in real time. Thus choosing a wavelet for eye detection depends on a lot of trial and error. Discrete Wavelet Transform is recursively applied to all the images in the training data set until the lowest frequency subband is of size 32x32 pixels i.e. the LH subband at a particular level or depth of DWT is of size 32x32. The original image's grayscale image is shown in figure 5.2. The LH subband at resolution 32x32 is shown in Figure 5.3.Here we have used HAAR wavelet instead of Gabor wavelet while calculating wavelet transform.

fig-5.2 the original image

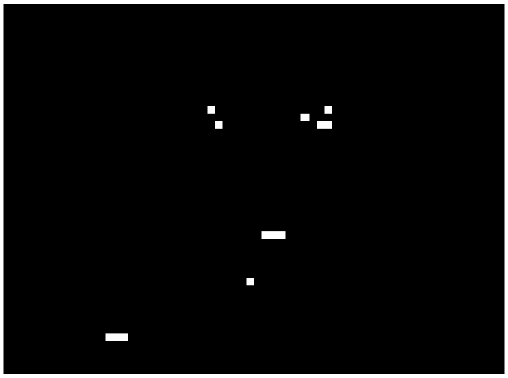
fig-5.3 the LH sub-band figure

We take the modulus of the wavelet coefficients in the LH subband. Experiments were performed to go to a resolution even coarser than 32x32. However, it was observed that in certain cases the features would be too close to each other and it was difficult even manually too to separate them. This would burden the Neural Network model and a small error in locating the eyes at this low resolution would result in a large error in locating the eyes in the original image.

5.4 Detection of Wavelet Maxima:

Our approach to eye detection is based on the observation that, in intensity images eyes differ from the rest of the face because of their low intensity. Even if the eyes are closed, the darkness of the eye sockets is sufficient to extract the eye regions. These intensity peaks are well captured by the wavelet coefficients. Thus, wavelet coefficients have a high value at the coordinates surrounding the eyes. We then detect the wavelet maxima or the wavelet peaks in this LH subband of resolution 32x32. Note that several such peaks are detected, which can be the potential locations of the eyes. The intensity peaks are shown in Figure 5.4 and 5.5.

Fig-5.4 the wavelet maxima of LH sub-band



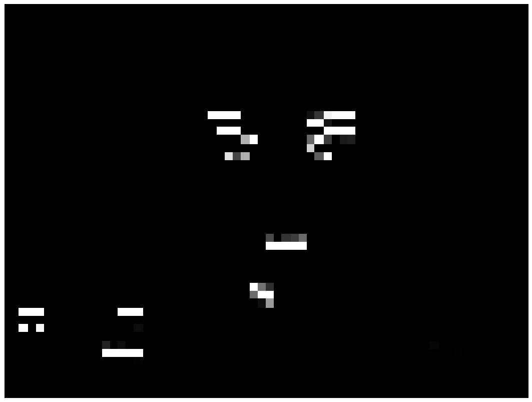
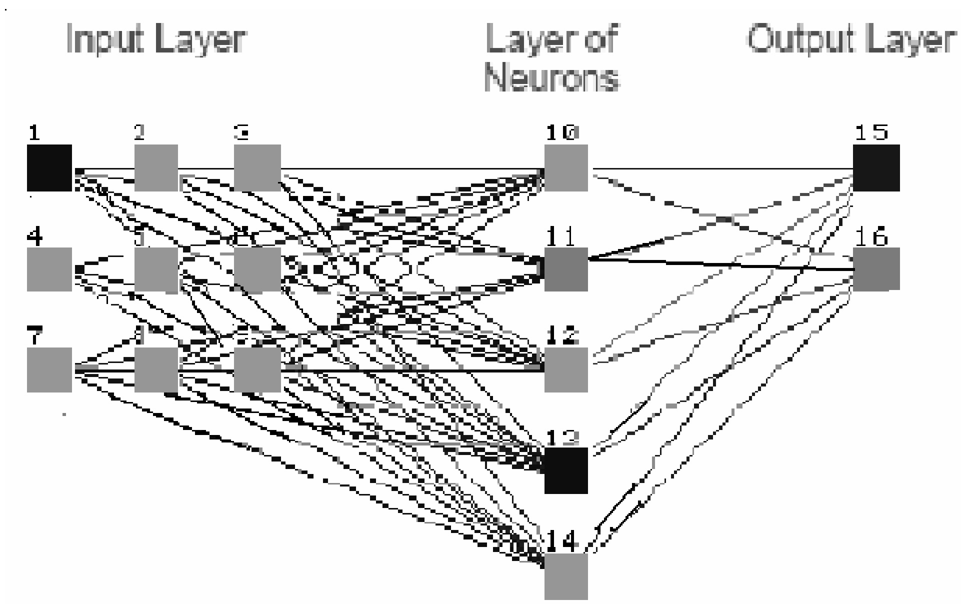


Fig-5.5 LH sub-band with peaks replaced by its 3*3 neighborhood wavelet coefficient

5.5 Neural Network Training:

The wavelet peaks detected are the center of potential eye windows. We then feed 3x3 neighborhood wavelet coefficients of each of these local maxima's in 32x32 LH subbands of all training images to a Neural Network for training. The Neural Network has 9 input nodes, 4 hidden nodes, and 2 output nodes. A diagram of the Neural Network architecture is shown in Figure 5.6.. A (1,-1) at the output of Neural Network indicates an eye at the location of the wavelet maxima whereas (-1, 1) indicates a non-eye. Two output nodes instead of one were taken to improve the performance of the Neural Network. MATLAB's Neural Network Toolbox was used for simulation of the back propagation Neural Network. A conjugate gradient learning rate of 0.4 was chosen while training. This completes the training stages for neural networks back propagation model.



Here we have used the MLP (multi-layer perceptions) back-propagation model for neural network training. It consists of having 9 neurons in the input layer. In the hidden layer or 2nd layer has 5 neuron for processing .In the 3rd or output layer has the two nodes for showing the output. We have taken two output node insists of one to get a better accuracy towards detecting eye. After this you have to find eye part & non-eye part in the figure from neural network model. Where an output of (1, -1) indicates the presence of an eye & output of (-1,1) indicates the presence of a non-eye.

Chapter 6

EXPERIMENTATION & RESULTS

7.1 SOURCE CODE

Program no.1

(Wavelet transform part)

```
 function[inp]=aisT(a)
% clear all;
% close all;
% clc;
dwtmode('zpd');
% a=double(rgb2gray(imread('C:\MATLAB7\work\project\ais2.jpg')));
 imshow(uint8(a))
 %figure
sizea=size(a)
[c,s]=  wavedec2(a,4,'haar');
            % ap_cf = appcoef2(c,s,'haar',4);
            % sizeAP_CF=size(ap_cf)
sizec=size(c)
h_cf2 = detcoef2('h',c,s,4);

[m,n]=size(h_cf2);      %n used later

sizehcf=size(h_cf2)
% sizevcf=size(v_cf2)
% sizedcf=size(d_cf2)

% imshow(uint8(ap_cf))             %for LL image
 %figure                           %figure of LH
```

```
%imshow(uint8(h_cf2))
imview(h_cf2)
% % figure
% % imshow(uint8(v_cf2))
% % figure
% % imshow(uint8(d_cf2))

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%          DETECTION  OF WAVELET
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%                    MAXIMA

% LH=abs(h_cf2);
LH1=zeros(m*n,1);           % LH1 just used for descending matrix h_cf....

%^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
k=1;
for i=1:m
    for j=1:n
        LH1(k)= abs(h_cf2(i,j));      %(h_cf2(i,j));
        k=k+1;
    end
end
   LH=sort(LH1,'descend');                    %all values h_cf2 r stored in LH1
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%max
peak wavelet
ln=25;
max_pk=zeros(ln,1);
for i=1:ln
    max_pk(i)=LH(i);                              %the value of top ln=25 max wave coef.. in LH
figure(h_cf2)
end
```

```
% % % % % % % % % % % % % % % % % % % % % % % % % % % % % % %

dhcf2=zeros(m,n);
for i=1:m
    for j=1:n
        if any(  abs(h_cf2(i,j))>=max_pk )     %  (h_cf2(i,j))>=max_pk )
            dhcf2(i,j)=h_cf2(i,j);                                  % dchf is used for max peak detectiopn
&                                                                    others zero
        end
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%
% dhcf2
imview(dhcf2)
%    figure
%    imshow(uint8(dhcf2))
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%        for
neighbourhood configuration
   inp=zeros(ln,13);
   size(inp)
   l=1;k=1;
   ph_cf2=[zeros(1,n+2);zeros(m,1),h_cf2,zeros(m,1);zeros(1,n+2)];        %2   avoid   error
Attempted to access h_cf2(13,51);
   pdhcf2=[zeros(1,n+2);zeros(m,1),dhcf2,zeros(m,1);zeros(1,n+2)];          %index  out  of
bounds because size(h_cf2)=[38,50].

   [size_LHx,size_LHy]=size(h_cf2);

   for i=2:m+1
      for j=2:n+1
```

```
        if dhcf2(i-1,j-1)~=0

            pdhcf2(i-1,j-1)=ph_cf2(i-1,j-1);                pdhcf2(i-1,j)=ph_cf2(i-1,j);       pdhcf2(i-
1,j+1)=ph_cf2(i-1,j+1);
            pdhcf2(i,j-1)=ph_cf2(i,j-1);                                               pdhcf2(i,j)=ph_cf2(i,j);
pdhcf2(i,j+1)=ph_cf2(i,j+1);
            pdhcf2(i+1,j-1)=ph_cf2(i+1,j-1);                                  pdhcf2(i+1,j)=ph_cf2(i+1,j);
pdhcf2(i+1,j+1)=ph_cf2(i+1,j+1);

              inp(l,k)   =ph_cf2(i-1,j-1);           inp(l,k+1)=ph_cf2(i-1,j);     inp(l,k+2)=ph_cf2(i-
1,j+1);
              inp(l,k+3)=ph_cf2(i,j-1);                                                  inp(l,k+4)=ph_cf2(i,j);
inp(l,k+5)=ph_cf2(i,j+1);
              inp(l,k+6)=ph_cf2(i+1,j-1);                                          inp(l,k+7)=ph_cf2(i+1,j);
inp(l,k+8)=ph_cf2(i+1,j+1);
%                 inp(l,k+10)=i-1/size_LHx;       inp(l,k+9)=j-1/size_LHy;
                   inp(l,k+10)=i-1;       inp(l,k+9)=j-1;
             l=l+1;     k=1;

        end
      end
   end
   inp(:,10:11)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
   for i=1:m
    for j=1:n
        xdhcf2(i,j)=pdhcf2(i+1,j+1);
    end
   end
   imview(xdhcf2)
%     figure
```

```
 %       imshow(uint8(xdhcf2))
% max_pk
% pdhcf2
% ph_cf2
```

Program no.2

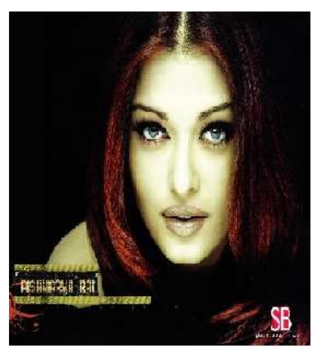
(Neural network model part)

```
%for aiswariya

close all;

clear all;

clc;

a=double(rgb2gray(imread('C:\MATLAB7\work\project\ais2.jpg')));

pa=aisX(a);

pk=pa(:,1:9);

p=pk';

 Tp=[1 1 1 1 1 1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1];

 T=[Tp;-Tp];

 net = newff( minmax(p),[5 2],{'tansig' 'purelin'});

 Y = sim(net,p);

%  plot(P,T,P,Y,'o');
```

```
%   net.trainParam.show = 50;

  net.trainParam.lr = 0.4;

  net.trainParam.epochs = 4000;

  net.trainParam.goal = 1e-5;

  [net,tr] = train(net,p,T);

  Y = sim(net,p)

%   plot(P,T,P,Y,'o')

%   Y=sim(net,3.5)

  b=double(rgb2gray(imread('C:\MATLAB7\work\project\ais6.jpg')));

  pb=aisT(b);

  pbc=pb(:,1:9);

  Y=sim(net,pbc')
```

Input: (1)

Original image for training neural network ------------→
(training image)



Input: (2)
2^{nd} pic of the same person for detection(test image)

Output :

Gray-scale image of input image-------------------------→

LH sub-band image-------------------------------------→

Wavelet maxima of LH sub-band-----------------------→

Sub-band with peak replaced by 3*3 neighborhood------→

Y1(output for training image) =

Columns 1 through 7

0.9090	0.9173	0.9187	0.8910	0.9654	0.9287	-0.8449
-0.9604	-0.9515	-0.8977	-1.1043	-0.9719	-0.9439	0.7588

Columns 8 through 14

-0.9669	-0.8692	-0.9752	-0.9505	-0.9779	-0.9481	-0.0072
0.8855	0.9132	0.7044	0.8979	0.7072	0.9427	0.8075

Columns 15 through 21

-0.9867	-0.8649	-0.9705	-0.9671	-0.9180	-0.8279	-0.9568
0.9640	0.9966	0.9099	0.8198	0.8876	0.9275	0.9610

Columns 22 through 25

-0.9707	-0.9030	-0.9872	-0.9412
0.9744	0.9312	0.9944	0.9187

Y2(output for test image)=

Columns 1 through 7

-0.9593	-0.8823	0.9363	0.9128	0.8900	0.9825	0.9067
0.8770	0.9303	-0.9011	-0.9900	-0.9881	-0.9614	- 0.9667

Columns 8 through 14

0.8744	-0.8321	-0.8557	-0.9644	-0.9089	-0.8669	-0.9586
-0.9106	0.9858	0.9368	0.9782	0.8298	0.9175	0.8018

Columns 15 through 21

-0.9604	-0.8081	-0.8315	-0.8783	-0.8116	-0.9650	-0.9011
0.8948	0.9015	0.9271	0.9845	0.9891	0.9533	0.8772

Columns 22 through 25

-0.8403	-0.8179	-0.9558	-0.9443
0.9556	0.9958	0.8170	1.0770

Chapter 7

CONCLUSION

7.1. Performance

A number of experiments were done to test the robustness of the algorithm and to increase the accuracy of eye detection. Various architectures of Neural Networks with different learning rates were tried and it was found that back propagation with conjugate gradient learning seemed to be the best choice. A very high learning rate of 0.8 was Chosen because the learning algorithm was getting trapped in local minima while training the network. Final training was stopped when the error graph, as shown in Figure 7.1, didn't show any significant fluctuation.

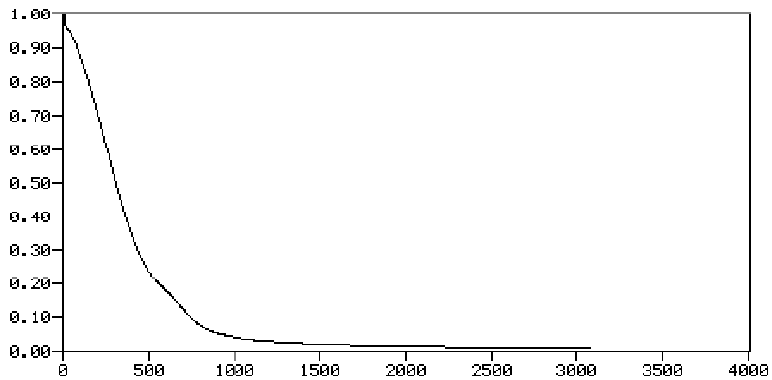


Figure 7.1- Conjugate Training Error Curve

An experiment was done in which the face was analyzed using wavelet packets and it was found that most of the information was retained by the low frequency sub bands and the high frequency packets had no information. Images with different states of the eye (closed, open, half open, looking sideways, head tilted etc.) and varying eye width were chosen. The eye positions found were compared with the positions that were pointed out manually. The eyes were correctly located when its location is within two pixels, in both x and y directions, of the manually assigned point. The variation of 2 pixels is deliberately allowed, to compensate for the inaccuracies in the location of eyes during training. An accuracy of 88% was observed in the final location of the eyes. A database of 60 test images was evaluated for performance. All these test images were captured in totally different environment conditions and were not included while training the Neural Network. Most of the error cases occurred in images with complex background. Also there was an error in accurately determining the exact location of the eyes since a 1-pixel shift at a resolution of 32x32 corresponded to a larger shift in the exact location of the eye. In some cases the Neural Network classified only 1 peak as an eye in spite of the presence of 2 eyes in the image. In a few cases observations were made in which regions of the face not belonging to the eyes were detected as eyes. In other cases more than 2 eyes were indicated in the image. In contrast, the performance of this algorithm, which uses wavelets as a preprocessor to Neural Networks, the algorithm with only Neural Networks, achieved an accuracy of 81% in detecting the exact location of the eyes.

7.2 CONCLUSION:

This type of approach gives a new dimension to the existing eye detection algorithms. The present algorithm is robust and at par with the other existing methods but still has a lot of scope for improvement. In this type of approach a wavelet subband approach in using Neural Networks for eye detection. Wavelet Transform is adopted to decompose an image into different subbands with different frequency components. A low frequency subband is selected for feature extraction. The proposed method is robust against illumination, background, facial expression changes and also works for images of different sizes. However, a combination of information in different frequency bands at different scales, or using multiple cues can even give better performance. Further studies in using Fuzzy Logic for data fusion of multiple cues will give better results.

REFERENCES:

1. M.J.T. Reinders, "Eye tracking by template matching using an automatic codebook generation scheme", Third annual conference of the Advanced School for Computing and Imaging, pp. 85-91, Heijen, The Netherlands, June 1997.

2. Kumar, Thilak R and Raja, Kumar S and Ramakrishnan, “Eye detection using color cues and projection functions”, Proceedings 2002 International Conference on Image Processing ICIP, pages Vol.3 337-340, Rochester, New York, USA.

3. K. M. Lam, H. Yan, ” Locating and extracting the eye in human face images”, Pattern Recognition, Vol. 29, No. 5 pp.771-779.(1996)

4. Takacs, B., Wechsler, H., "Detection of faces and facial landmarks using iconic filter banks", Pattern Recognition, Vol. 30, No. 10, October 1997, pp. 1623-1636.

5. Vezhnevets V., Degtiareva A., "Robust and Accurate Eye Contour Extraction", Proc. Graphicon-2003, pp. 81-84, Moscow, Russia, September 2003.

6. Erik Hjelms and Jrn Wroldsen, "Recognizing faces from the eyes only", Proceedings of the 11th Scandinavian Conference on Image Analysis, 1999

7. A. Pentland, B. Moghaddam, T. Starner,“View-based and modular eigenspaces for face recognition”, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Seattle, WA, 1994, pp.84-91.

8. C. Morimoto, D. Koons, A. Amir, and M.Flickner, “Real-Time Detection of Eyes and Faces”, Proceedings of 1998 Workshop on Perceptual User Interfaces, pages 117-120, San Francisco, CA, November 1998.

9. W. Sweldens and P. Schrder, "Building your own wavelets at home", Wavelets in Computer Graphics, pp. 15--87, ACM SIGGRAPH Course notes, 1996.

10.Baback Moghaddam and Ming-Hsuan Yang. Gender Classification with Support Vector Machine, Proceeding of the 4th International Conference on Face and Gesture Recognition, pp306-311, Grenoble, France, 2000.

11. L. Ma, Y. Wang, T. Tan. Iris recognition using circular symmetric filters. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 2002.

12. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, pp. 3445-3463, Dec. 1993.